The Report Committee for Megan Michelle Moeller
Certifies that this is the approved version of the following report:

Methods for Analyzing Proportions

APPROVED BY
SUPERVISING COMMITTEE:

Supervisor:

Stephen Jessee

Michael Scott Wolford

Methods for Analyzing Proportions

by

Megan Michelle Moeller, B.A.

Report

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Science in Statistics

The University of Texas at Austin

August 2013

Dedication

This master's report is dedicated to my loving and supportive partner, Andy.

Acknowledgements

I owe my gratitude to my Committee, Dr. Jessee and Dr. Wolford, for their time and effort on this project. I am indebted to Dr. Sean Theriault for his unwavering support and patience throughout the last four years. Finally, I cannot thank Dr. Bob Luskin enough for piquing my interest in Statistics and for being such a phenomenal teacher of the subject.

Abstract

Methods for Analyzing Proportions

Megan Michelle Moeller, M.S., Stat.

The University of Texas at Austin, 2013

Supervisor: Stephen Jessee

The analysis of proportions is interesting and noteworthy in that there are no commonly accepted regression models for analyzing proportions; indeed, researchers most often use ordinary least squares to estimate the parameters of a linear regression model for proportional data. Such an approach, however, violates several assumptions of the Classical Linear Regression Model. This report outlines the general linear model and the problems associated with using this approach to model proportions and considers a variety of alternate approaches that researchers have taken to model proportions. These alternatives include transforming the dependent variable, a censored regression (Tobit) model, a Fractional Logit model, and Beta Regression. All of the approaches considered are implemented in a case study analyzing Rice party difference scores in the 93rd to 108th Congress. A comparison of the results from each approach confirms the findings of other researchers that Beta regression is the most preferred approach for modeling proportions.

Table of Contents

List of Tables

List of Figures

List of Illustrations

Chapter 1: Introduction

In Political Science research, scholars often analyze a proportion or some other measure bounded from 0 to 1 as the dependent variable. Due to the fact that a proportion cannot take on any value outside of this interval, variables in the form of a proportion are a type of Limited Dependent Variables. Limited Dependent Variables are those "for which observations are limited to a certain range" (Kmenta 1997: 560). Variables in the form of proportions are not the only type of Limited Dependent Variable. Other instances in which the range of values the dependent variable can take on is limited include when the dependent variable is *censored* or *truncated*. Censoring occurs when certain values of the dependent variable are unobservable, resulting in missing information. Truncation occurs in samples in which not only the values of the dependent variable are unobservable, but also, as a result, the values of the corresponding independent variables are unobservable and the number of missing observations is unknown (Kmenta 1997). Dependent variables with discrete outcomes, such as binary variables, some types of ordinal variables, and count variables, are also limited dependent variables in that the values that those types of variables can take on are restricted in some way (Long 1997).

Studies that examine how a variety of independent variables influence a proportion or vector of proportions can be divided into two categories. The first category analyzes single proportions and the second category analyzes multiple proportions (Buis 2010). This report, however, only focuses on the first category of single proportions.

Analysis of proportional data is interesting and noteworthy in that there are no commonly accepted regression models for analyzing proportions; indeed, researchers most often use ordinary least squares to estimate the parameters of a linear regression model for proportional data (Kieschnick and McCullough 2003). As this report will demonstrate, such an approach violates several assumptions of the Classical Linear Regression Model.

This report will proceed by first outlining the general linear model and the problems associated with using this approach to model proportions. Second, this report will consider a variety of alternate approaches that researchers have taken to model proportions. These approaches can be divided into four broad categories: 1) transforming the proportion such that the response variable will be able to take on any real value before performing least squares estimation; 2) a censored regression model estimated by maximum likelihood in which the researcher erroneously assumes that the dependent variable is a latent variable normally distributed over $\mathbb{R}$, but is only observed over the interval [0, 1]; 3) a quasi-parametric Generalized Linear Model which models how only the mean proportion relates to explanatory variables; 4) Beta Regression, which implements a Generalized Linear Model assuming the dependent variable follows a beta distribution. Last, this report will implement these methods using a case study and compare the results.

Chapter 2: Modeling Proportions with the General Linear Model

Proportions data in a variety of fields is most commonly modeled using the general linear model and ordinary least squares estimation (Kieschnick and McCullough 2003). This report will first provide an overview of the general linear model and the associated assumptions and will then discuss some of the concerns regarding the use of this model with proportions data.

The General Linear Model

The general linear model is given by the equation:

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \varepsilon_i \qquad i = 1, 2, \ldots, n \qquad (2.1)$$

For a sample with n observations, there will be n such equations, which can be written in matrix terms as:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \qquad (2.2)$$

where $\mathbf{Y}$ is an n x 1 vector of responses, $\mathbf{X}$ is an n x k+1 model matrix with columns for regressors, $\boldsymbol{\beta}$ is a k+1 x 1 vector of regression coefficients, and $\boldsymbol{\varepsilon}$ is an n x 1 vector or errors. The least squares estimators and the maximum-likelihood estimators for the multivariate linear model are equivalent, and are given by:

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \qquad (2.3)$$

The full specification of the "classical normal linear regression model" not only includes the form of the regression equation in (1.1) and (1.2), but also the following key assumptions regarding the specification of the probability distribution of the disturbance,

or the error term ε_i (Kmenta 1997). First, the expectation, or the mean, of ε_i is 0. That is, $E(\varepsilon_i) = 0$. This allows the expectation of the dependent variable to be a linear function of the explanatory variable.[1] Second, ε_i is homoskedastic. That is, $\text{Var}(\varepsilon_i) = \sigma^2$, meaning every disturbance has the same, constant variance σ^2 whose value is unknown. Third, the disturbances are uncorrelated, that is any ε_i and ε_j are independent for $i \neq j$. This assumption of nonautocorreltion can be written $\text{Cov}(\varepsilon_i, \varepsilon_j) = 0$ where $i \neq j$ and will likely be met if the data are sampled independently from a large population. Fourth, ε_i is normally distributed. Specifically, $\varepsilon_i \sim N(0, \sigma^2)$ as given by the prior assumptions. Last, the elements in $\mathbf{X}$ are non-stochastic with values fixed in repeated samples, and the matrix $(1/n)(\mathbf{X}'\mathbf{X})$ is nonsingular and its elements are finite as n approaches infinity. One implication of this assumption is that $E(\varepsilon_i X_j) = X_j E(\varepsilon_i) = 0$ for all i, j (Fox 2008; Kmenta 1997). In matrix notation, these assumptions can be compactly written $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{I})$ and imply $\mathbf{Y} \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$.

Problems with Modeling Proportions Using the General Linear Model

When the dependent variable is a proportion, several of the key classical linear regression assumptions are violated. The first problem with using ordinary least squares to model proportions is that the effect of the explanatory variables tends to be nonlinear since the conditional expectation function maps onto a bounded interval. It is clear that proportions are not normally distributed since they are only defined in the interval (0, 1), and the normal distribution is defined over $\mathbb{R}$, the set of all real numbers. The general

[1] $\boldsymbol{\mu} \equiv E(\mathbf{Y}) = E(\boldsymbol{Y}|\boldsymbol{X}) = E(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}) = \mathbf{X}\boldsymbol{\beta} + E(\boldsymbol{\varepsilon}) = \mathbf{X}\boldsymbol{\beta}$

linear model assumes a conditional normal distribution for the model and that the conditional expectation function is linear—that is, $E(\mathbf{Y}|\mathbf{X}) = \mathbf{X\beta}$ in matrix notation for n equations. This is problematic for modeling proportions because the linear model cannot confine predicted values of Y_i to the interval (0,1). The fact that proportions are only observed over a closed interval implies that the conditional expectation function must be non-linear (Kieschnick and McCullough 2003). The troubling implication of nonlinearity in a general linear regression model is that "the model fails to capture the systematic pattern of relationship between the response and the explanatory variables," and the results of estimation can be severely misleading (Fox 2008: 227).[2] For example, as previously mentioned, ordinary least squares estimation might generate impossible predicted values of the response variable, such as negative values or values greater than 1. The fact that a model to predict proportions could generate predictions that are not proportions provides a strong argument for departing from the general linear model.

Another complication that arises when modeling proportions with general linear models is that the error variance tends to be heteroskedastic, approaching 0 as the mean gets closer to either of the boundaries (Kieschnick and McCullough 2003). Constraints on the disturbance, which follow directly from constraints on the range of the dependent variable, also imply that the disturbance is data dependent. The variance, therefore, will not be constant for all observations (Greene 2008). Heteroskedasticity not only impairs the efficiency of the least-squares estimator, but also causes the standard error estimates to be biased and inconsistent, leading to incorrect confidence intervals, and ultimately

[2] In the case that the misspecification of the model is due to an omitted relevant explanatory variable, the

undermining inferences about the population coefficients. Although it is possible to estimate a heteroskedastic OLS model, failing to model the variance as a function of the independent variables may lead researchers to overlook theoretically interesting aspects of the data generating process (Paolino 2001). A third problem with using ordinary least squares on proportions data is that measures of proportions often display asymmetry, and the errors tend not to be normally distributed (Buis 2006; Ferrari and Cribari-Neto 2004). Although the violation of the assumption of normality seems relatively innocuous to least squares estimation because of the central-limit theorem, if the error distribution is highly skewed, it can compromise the interpretation of the least-squares fit (Fox 2008; Kmenta 1997).

Chapter 3: Alternate Approaches to Modeling Proportions

Transformations of the Dependent Variable

One of the most common methods researchers use to surmount some of these problems is to transform the dependent variable. Once the proportion has been transformed, researches will often proceed implementing ordinary least squares on the transformed response. The most frequent practice is to perform the logit transformation on the dependent variable (Kieschnick and McCullough 2003; Paolino 2001). The logit transformation,

$$y^* = \text{logit}(y) = \ln\left(\frac{y}{1-y}\right) \tag{3.1}$$

is helpful because it removes the upper and lower bounds of the scale, spreads out the tails of the distribution, and makes the resulting values symmetric about 0 (Fox 2008). In this approach, researchers estimate:

$$\mathbf{Y}^* = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{3.2}$$

where $\mathbf{Y^*}$ represents the transformed response variable. This method remedies the possibility of generating non-proportion expected values since the expectation can easily be transformed back to the original scale:

$$y = \frac{\exp(y^*)}{1+\exp(y^*)} \tag{3.3}$$

It has been shown that if y truly follows a logit-normal distribution with probability density function,

$$f_X(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\text{logit}(x)-\mu)^2}{2\sigma^2}} \frac{1}{x(1-x)},\ x \in (0,1) \tag{3.4}$$

then y* will follow a normal distribution, $N(\mu, \sigma^2)$ and ε_i will follow a standard normal distribution (Aitchison 1986). The assumption that y follows a logit-normal distribution can be verified by testing whether ε follows a standard normal distribution (Kieschnick and McCullough 2003).

A less frequently used transformation for modeling proportions is the probit transformation,

$$y^* = \text{probit}(y) = \Phi^{-1}(y) \tag{3.5}$$

where Φ^{-1} is the inverse cumulative distribution function for the standard normal distribution.[3] The inverse of the cumulative distribution function (CDF) can be used here because the range of the CDF is [0, 1].[4] The mapping of the response variable to a normal CDF produces a transformation very similar to that of the logit transformation. In fact, logit ≈ (π/3)·probit when their scales are equated (Fox 2008). One difference between the two is that the logit transformation has heavier tails compared to the probit, but this consideration is likely only relevant when the researcher is particularly interested in the proportions close to the tails (Hox 2010). Neither the probit or logit transformations can be applied to proportions that are exactly 0 or exactly 1. In the case that the response variable does take on the values of 0 or 1, the untransformed proportions can be mapped onto an interval that does not contain 0 or 1, and the logit transformation will be performed on the result. For example, $P' = .025 + (.097 \cdot P)$ maps the proportion to the interval [.025, .097] where P is the original proportion. This way,

[3] The inverse cumulative distribution function is also known as the quantile function. The quantile function is well known in statistical analysis for its ability to diagnose deviations from normality via Q-Q plots.
[4] The range of any continuous CDF is [0, 1].

when logit(P') transformation is applied, even proportions that were initially 0 or 1 are transformed. One drawback to this method is that neither transformation will remedy the violation of homoskedasticity (Kieschnick and McCullough 2003).[5]

Tobit Model

The Tobit Model, also called the *censored regression model*, is another model that researchers have used to model proportional data. The original censored regression model, proposed by James Tobin in 1958, was designed for data censored to the left at 0 and allowed the upper bound of the data to extend to infinity. The model assumes that there is some latent response variable ξ_i whose values cannot be observed outside of a certain range (a, b). The latent variable ξ_i is assumed to be normally distributed, such that $N(\mu, \sigma^2)$, is assumed to be linearly related to the explanatory variables, such that

$$\xi_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \varepsilon_i \tag{3.6}$$

and all of the other assumptions regarding the disturbance term from the normal regression model hold as well (Fox 2008). In matrix notation, the regression equation can be written,

$$\boldsymbol{\xi} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{3.7}$$

Note that the dimensions of these matrices are equivalent to those defined in (2.2). To analyze censored data, a new random variable y_i^*, transformed from the latent variable ξ_i, is defined:

[5] A less common, more archaic alternative transformation is the arcsine-square-root transformation which is a variance stabilizing transformation (Fox 2008).

$$y_i^* = \begin{cases} a & \text{for } \xi_i \le a \\ \xi_i & \text{for } a \le \xi_i \le b \\ b & \text{for } \xi_i \ge b \end{cases} \quad (3.8)$$

When $a = 0$ and $b = \infty$, the probability distribution that applies to the observed y_i^* is a mixture of discrete and continuous, given by $\Pr(y_i^* = 0) = \Pr(\xi_i \le 0) = \Phi(-\mu/\sigma) = 1 - \Phi(\mu/\sigma)$, and if $\xi_i > 0$, then y_i^* has the density of ξ_i (Greene 2008). Tobit models are typically estimated using the method of maximum likelihood estimation. This because maximum likelihood is the method best suited to estimate models in which the parameters are related in a nonlinear way to the mean response function, as in (3.12). For the case described above where $a = 0$ and $b = \infty$, the log-likelihood for censored regression model is

$$\ln L = \sum_{Y^*=0} \ln\left[\, 1 - \Phi\left(\frac{\mathbf{X}\boldsymbol{\beta}}{\sigma}\right)\right] + \sum_{Y^*>0} -\frac{1}{2}\left[\ln 2\pi + \ln\sigma^2 + \frac{(Y^* - \mathbf{X}\boldsymbol{\beta})^2}{\sigma^2}\right] \quad (3.9)$$

with marginal effects,

$$\frac{\partial E[\boldsymbol{Y}^*|\mathbf{X}]}{\partial \mathbf{X}} = \boldsymbol{\beta} \times \Pr(a < Y^* < b) \text{ (Greene 2008)} \quad (3.10)$$

When researchers use the Tobit model on proportional data, they assume everything previously stated except that ξ_i is censored from (0, 1), so the observed response variable y_i^* is defined by,

$$y_i^* = \begin{matrix} 0 & \text{for } \xi_i \le 0 \\ \xi_i & \text{for } 0 \le \xi_i \le 1 \\ 1 & \text{for } \xi_i \ge 1 \end{matrix} \quad (3.11)$$

The conditional expectation function for the observed data is,

$$E(\boldsymbol{Y}^*|\mathbf{X}) = \Phi\left(\frac{\mathbf{X}\boldsymbol{\beta}}{\sigma}\right)(\mathbf{X}\boldsymbol{\beta} + \sigma\lambda) \quad \text{where } \lambda = \frac{\phi(\frac{\mathbf{X}\boldsymbol{\beta}}{\sigma})}{\Phi(\frac{\mathbf{X}\boldsymbol{\beta}}{\sigma})} \quad (3.12)$$

The log-likelihood function in this case is

$$\ln L = \sum_{Y^*=0} \ln\left[1 - \Phi\left(\frac{\mathbf{X}\boldsymbol{\beta}}{\sigma}\right)\right] + \sum_{0<Y^*<1} \ln\left[\frac{1}{\sigma}\phi\left(\frac{Y^* - \mathbf{X}\boldsymbol{\beta}}{\sigma}\right)\right]$$
$$+ \sum_{Y^*=1} \ln\left[\Phi\left(\frac{\mathbf{X}\boldsymbol{\beta}-1}{\sigma}\right)\right] \quad (3.13)$$

One of the primary problems with using the Tobit method is that for the data observed on the interval (0, 1) the regression results will be observationally equivalent to the general linear model, and so all the same criticisms from that model apply to the Tobit as well. Another, more theoretical, problem with Tobit is that it assumes ξ_i is normally distributed, but only observes values within the range [0, 1]. Proportions data provides observations from [0, 1] not because the variable of interest is censored, but rather because proportions are not defined outside this interval (Kieschnick and McCullough 2003). Using a censored regression model to model non-censored data is inherently troubling. Furthermore, this approach also fails remedy the violation of homoskedasticity (Kieschnick and McCullough 2003). In addition to impairing the ordinary least squares estimator, the presence of heteroskedasticity can cause very serious problems for the consistency of the maximum likelihood estimator (Greene 2008).[6] Greene also notes that if the disturbances in a Tobit model are not normally distributed, then the maximum likelihood estimator is inconsistent (Greene 2008).

[6] The degree to which heteroskedasticity impairs the consistency of the maximum likelihood estimator depends on the degree of censoring (Greene 2008).

Fractional Logit Generalized Linear Model

The Generalized Linear Model (GLM)[7] is a generalization of linear regression that allows a linear model to be related to a response variable that follows a non-normal distribution, such as a distribution that generates proportion values, constrained to the interval [0,1] (Fox 2008). A key component of a GLM is a smooth and invertible linearizing *link function* $g(\cdot)$ that transforms the mean of the response variable to the linear function of regressors. Recall that $E(Y) = \mu$.

$$g(\mu) = \eta = \boldsymbol{X\beta} \tag{3.14}$$

where η is called the *linear predictor* – that is, the linear model. The expected value of the response variable, μ, is linked to the linear predictor by the link function. Furthermore, because the link function is invertible,

$$\mu = g^{-1}(\eta) \tag{3.15}$$

In the GLM framework, the conditional mean function of non-normally distributed random variables can be written in terms of a linear predictor and inverted link function as,

$$E(y_i \mid \mathbf{X}) = \mu = g^{-1}(\eta) = g^{-1}(\mathbf{X\beta}) \tag{3.16}$$

and, therefore, the model is essentially a linear model for a transformation of the expected response (Fox 2008). GLMs are fit to the data using the method of maximum likelihood estimation.

[7] Not to be confused with the general linear model.

Papke and Wooldridge develop a GLM for proportions data using quasi-likelihood methods that others have termed "Fractional Logit" (Buis 2010; Papke and Wooldridge 1996). They assume that, for all $i = 1, 2, \ldots, N$,[8]

$$E(y_i \mid \mathbf{x}_i) = g^{-1}(\mathbf{x}\boldsymbol{\beta}) \tag{3.17}$$

where $0 < y_i < 1$ and where $g^{-1}(z)$ is a known function satisfying $0 < g^{-1}(z) < 1$ for all $z \in \mathbb{R}$. Under these conditions, the predicted values of y will lie in the interval $(0, 1)$, and the conditional expectation function will be defined even if y takes on a value of exactly 0 or 1 (Papke and Wooldridge 1996). Note that no other assumption is made about the underlying structure that generates y_i. In this quasi-parametric approach, they do not explicitly model the probability distribution of the dependent variable; instead they simply assume a restricted range for the dependent variable and model how the mean proportion relates to the explanatory variables. Papke and Wooldridge, as well as others implementing this model, choose the logistic function (i.e. the inverse-logit function) for $g^{-1}(\cdot)$, such that $g^{-1}(z) = \frac{\exp(z)}{1+\exp(z)} \equiv \frac{1}{1+\exp(-z)}$, and so the link function $g(\cdot)$ is the logit link function given by (3.1), although the cumulative distribution function—that is, the inverse of the probit function—would be appropriate as well (Kieschnick and McCullough 2003; Papke and Wooldridge 1996). The Bernoulli log-likelihood function, given by,

$$\ln L = y_i \ln[g^{-1}(\mathbf{x}_i\boldsymbol{\beta})] + (1 - y_i) \ln[1 - g^{-1}(\mathbf{x}_i\boldsymbol{\beta})] \tag{3.18}$$

[8] Observations are assumed to be independent but not necessarily identically distributed. The methods in Papke and Wooldridge (1996) also ignore the information on n since the fraction y may not be a proportion from a discrete group size.

which is defined for $0 < g^{-1}(z) < 1$ for all $z \in \mathbb{R}$. Although the distribution of y_i is not specified, Papke and Wooldridge choose the Bernoulli log-likelihood because the Bernoulli is a member of the exponential family, because the log-likelihood function is easy to maximize, and because the quasi-maximum likelihood estimator of $\boldsymbol{\beta}$ is consistent, regardless of the true distribution of y_i (Papke and Wooldridge 1996). A quasi-maximum likelihood procedure is then used to estimate the parameters.

Beta Regression

In contrast to the quasi-parametric approach of the Fractional Logit GLM, another approach is to model the distribution of the dependent variable as a beta distribution. The beta distribution is a continuous distribution defined on the interval [0, 1], and its probability density function is given by,

$$f_X(x;\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \qquad 0 < x < 1 \tag{3.19}$$

where α and β are both positive-valued shape parameters. One of the most attractive features of the beta distribution is its versatility. The density of the beta distribution can take on many different shapes depending on the values of the shape parameters. Illustration 1 demonstrates a few of the many possible shapes the density function can take.

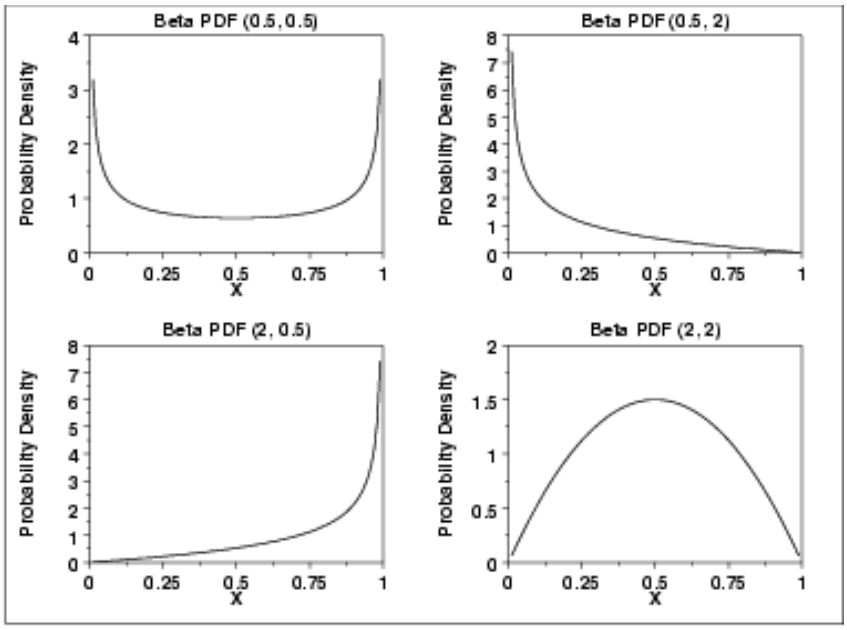

Illustration 1: Various Beta Probability Density Functions ("NIST/SEMATECH e-Handbook of Statistical Methods" n.d.).

The support and versatility of the beta distribution, therefore, makes it an attractive option for modeling proportions. The mean of a beta distribution is given by,

$$E(x) = \frac{\alpha}{\alpha+\beta} \tag{3.20}$$

The drawback of modeling proportions using this conventional parameterization of the beta distribution is that parameter estimates are difficult to interpret. In regression analysis, it is rarely interesting to estimate an explanatory variable's relationship to a shape parameter. It is much more desirable to model how the mean of the distribution of the dependent variable changes as the independent variables change (Buis 2006).

As a solution to this problem, some scholars have use an alternative parameterization of the beta distribution that defines a location parameter equivalent to the mean of the response. From (3.18), a new parameter μ is defined, such that,

$$E(y|\mathbf{x}) = \frac{\alpha}{\alpha+\beta} = \mu \quad \text{where } 0 < \mu < 1 \tag{3.21}$$

and define a precision parameter $\phi = \alpha+\beta$, where $\phi > 0$, so that $\alpha = \mu\phi$ and $\beta = (1-\mu)\phi$.[9] Using this alternative parameterization, the probability density function of a beta distributed random variable can be written,

$$f_y(y;\mu,\phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1}(1-y)^{(1-\mu)\phi-1}, \qquad 0 < y < 1 \tag{3.22}$$

and a beta-distributed random variable can now be modelled according to the structure of Generalized Linear Models specified in equation (3.15). As in the Factional Logit model, either the logit link or the probit link are appropriate choices for link functions since the response variable is defined on the interval (0, 1), but most researchers who implement this method choose the logit link (Buis 2010; Ferrari and Cribari-Neto 2004; Kieschnick and McCullough 2003; Paolino 2001). With this link function, the linear predictor can be written,

$$\eta = g(\mu) = \ln\left(\frac{\mu}{1-\mu}\right) = \boldsymbol{X\beta} \tag{3.23}$$

implying the mean of *y* in terms of the new parameterization can therefore be written,

$$E(y|\boldsymbol{x}) = \mu = g^{-1} = \frac{\exp(\mathbf{x}\boldsymbol{\beta})}{1+\exp(\mathbf{x}\boldsymbol{\beta})} \tag{3.24}$$

The log-likelihood function for the reparametarized beta distribution is then,

$$\ln L = \log\Gamma(\phi) - \log\Gamma(\mu\phi) - \log\Gamma((1-\mu)\phi) + (\mu\phi - 1)\log y + \big((1-\mu)\phi - 1\big)\log\Gamma(1-y) \tag{3.25}$$

[9] The variance, $Var(y) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$ in the conventional parameterization, will become $Var(y) = \mu(1-\mu)\frac{1}{1+\phi}$ in the alternative parameterization. ϕ can be interpreted as a precision parameter in that, when μ is held constant, as ϕ increases the variance of *y* decreases (Ferrari and Cribari-Neto 2004).

with μ defined so that (3.22) holds. Estimates for β are derived by maximizing the log-likelihood function with respect to the parameters β and ϕ. These estimates will retain all the desired properties of maximum likelihood estimators because the beta distribution is a member of the exponential family (Fox 2008). One caveat of Beta Regression is that it cannot handle proportions that are exactly 0 or 1, however it is possible to employ the $P' = .025 + (.097 \cdot P)$ transformation on the dependent variable to push the extreme values slightly inwards (Buis 2006).[10]

[10] Alternatively, a zero-one inflated beta model can be used in the case that 0 and 1 observations occur as a result of processes distinct from that which generates all of other observation (Buis 2010).

Chapter 4: A Case Study on Party Difference Scores

Within the context of legislative studies, one particular measure bounded from 0 to 1 is the Party Difference Score. Developed by sociologist Stuart Rice in the 1920s, it is also known as the Rice Difference Score and is calculated by taking the absolute difference between the proportion of Republicans and Democrats voting the same way on a given roll call vote (Rice 1925). This index is, in essence, a measure of inter-group difference or dissimilarity. Because it quantifies how unalike two parties are on a particular roll call vote or set of votes, legislative researchers commonly use the party difference score as a measure of party polarization and party conflict in the U.S. Congress (Lee 2009; Theriault 2008). For example, if every legislator votes the same way on a given vote, the party difference score would be 0. Further, if the same proportion of legislators from each party vote the same way, that is, if the extent of disagreement within one party on a vote is identical to the extent of disagreement within the opposing party, the party difference score for that vote will be 0. When a vote is equally divisive of both parties, the party difference score of 0 indicates that the two parties' preferences regarding that vote are not at all different from each other. Higher values of the party difference score indicate greater degrees of dissimilarity in the parties' preferences on given roll calls. A party difference score of 1 would reveal a strict party-line vote in which every member of one party votes Yea and every member of the opposing party votes Nay. Such a vote, in which every Democrat votes against every Republican, would indicate complete dissimilarity in the two parties' preferences on that roll call.

The party difference measure, one measure of party polarization, has played an important role in recent legislative studies because of the dramatic increase in polarization in the U.S. Congress since the 1970s. Figure 1 illustrates the rise in party conflict in the House of Representatives from the 93rd Congress (1973-1975) to the 112th Congress (2010-2012) measured by the mean party difference score for all roll call votes for each Congress.

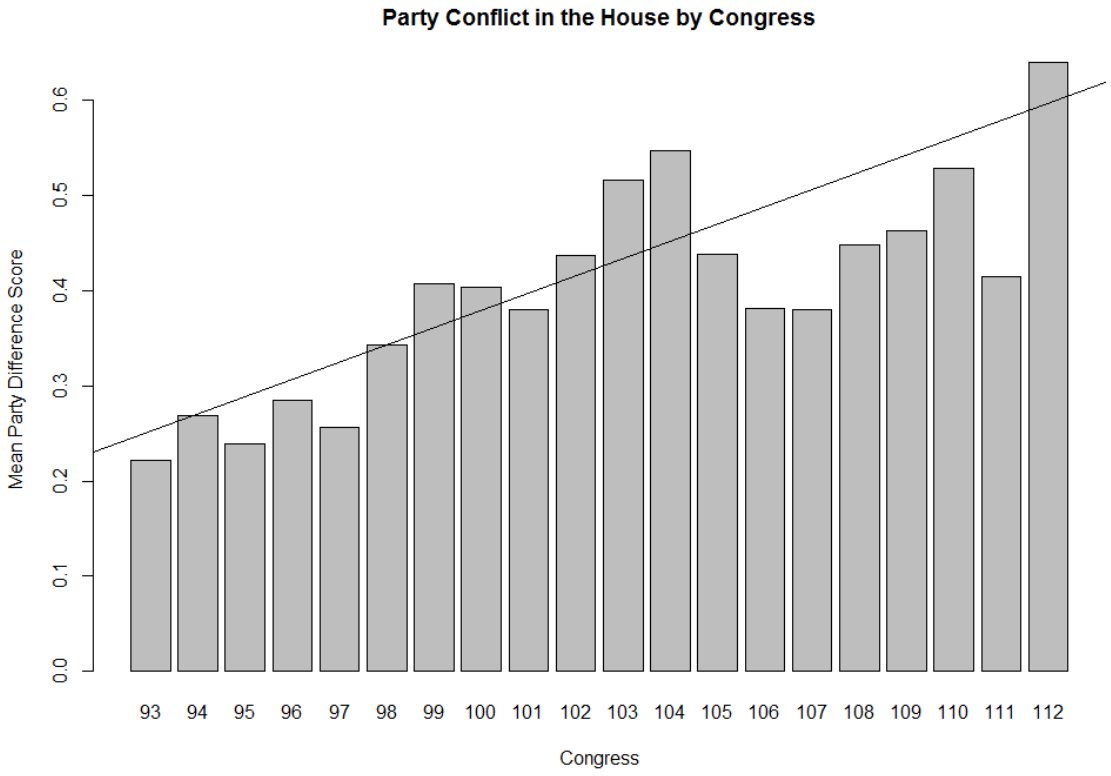

Figure 1.

The fitted line in the figure represents the results of simple regression analysis in which the mean party difference score for each congress is regressed on a time trend. The ordinary least squares coefficient estimate for time is 0.015, and is statistically significant at the 0.001 level, indicating that, on average, the parties became 1.5 percentage points

more divided on roll call votes each congress. Some scholars have shown that different types of votes (e.g. procedural votes) and votes on different issues (e.g. economic issues, social issues, etc.) produce varying mean party difference scores (Lee 2009; Theriault 2008). Figures 2 shows mean party differences scores in the House for the 93rd to 108th Congresses by issue. For further discussion on how votes are categorized into issue types, see Appendix A.

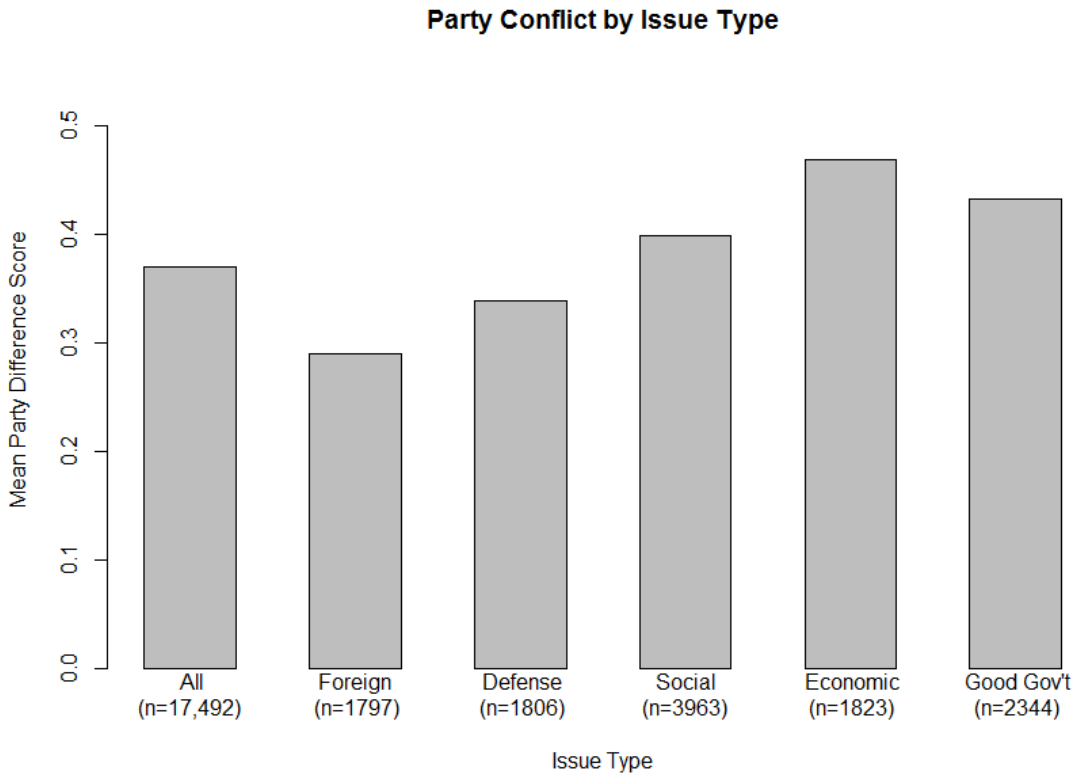

Figure 2.

Description of the data

In a classic work, one of the founders of the behavioral revolution in political science distinguishes between *position issues*, on which candidates and parties take

positions along a left-right ideological spectrum, and *valence issues*, on which everyone holds the same position (Stokes 1963). One example of a valence issue is government ethics; everyone is in favor of ethical behavior and no one is pro-corruption. *Good Government* issues are a set of valence issues defined as efforts to improve the government's integrity, efficiency, fairness, democratic accountability, and fiscal responsibility (Lee 2009). Both liberals and conservatives support virtue in government, so the position of favoring integrity, efficiency, and fairness in government cannot be placed on an ideological, left-right continuum. Good government issues can, nevertheless, be a source of conflict between political parties in Congress.[11] In order to analyze whether votes on certain issues, such as good government issues, display higher levels of partisanship than others, researchers will employ multivariate regression on party difference scores for each vote (Lee 2009). Lee shows that, controlling for issue, vote type, routine matters, and the factors specific to a given Congress, votes on good government issues are considerably more partisan than on the average vote in the Senate. Her conclusion comes from the fact that good government issues take a positive, statistically significant coefficient, indicating that, holding other factors constant, parties disagree more on good government votes than on the average vote in Congress. As previously mentioned, most scholars, including Lee, simply model party difference scores with a general linear model and ordinary least squares estimation. This report will

[11] For a full discussion of how and why valence issues can generate party conflict, see (Lee 2009).

conduct a similar analysis on votes in the House, implementing each approach described in Chapters 2 and 3, and will compares the results from each approach.[12]

Comparison of the Results

The first approach models the party difference scores using a general linear model. Figure 3 presents the Normal Q-Q Plot for the General Linear Model residuals

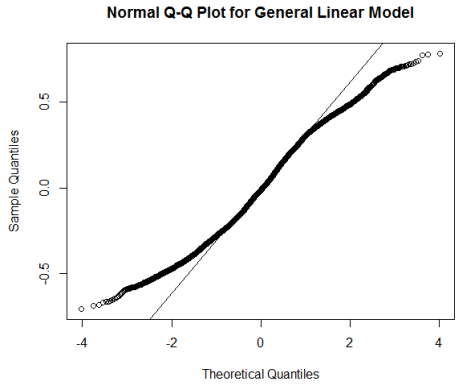

Figure 3.

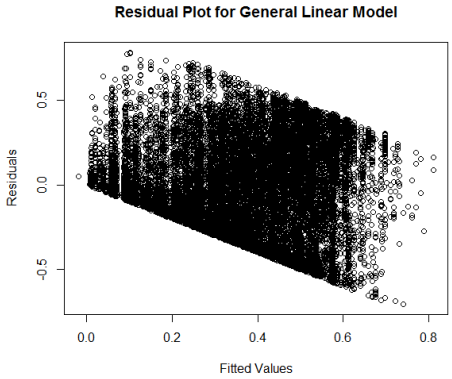

Figure 4.

and clearly indicates extreme departures from normality. More formally, the Shapiro-Wilk test rejects the normality of the residuals at 0.001 significance level. Figure 4 plots the residuals against the fitted values. Note that the manner in which the residuals deviate from 0 depends on the fitted values, implying a lack of linearity in the regression function and heteroskedasticity. The Breusch-Pagan test for non-constant error variance rejects the homoskedasticity of the residuals at 0.001 significance level. A summary of the fitted values also reveals that some fitted values do indeed fall outside of the [0, 1]

[12] This report will also control for whether or not the President has taken a position on the matter under consideration.

interval. It is, therefore, clear that the general linear model is an inappropriate choice to model party difference scores.

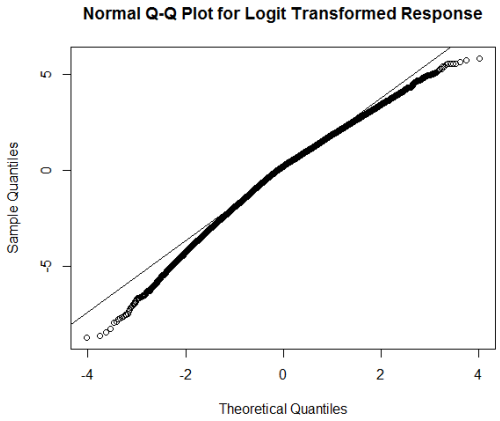

Figure 5.

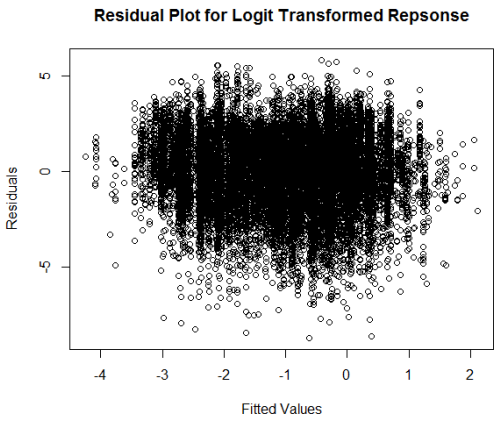

Figure 6.

Figure 5 presents the Normal Q-Q Plot for the residuals of the ordinary least squares model in which the response variable has been subject to a logit transformation. Although the Q-Q Plot indicates that the transformation has improved the normality, the Shapiro-Wilk test rejects the normality of the residuals at 0.001 significance level. Figure 6 plots the residuals against the fitted values. The result is an improvement upon that of the original general linear model. The residuals, however, seem to exhibit more deviation from 0 in the extremes of the fitted values than they do in the center. The results of the residual plot, however, may be more attributable to the covariates and their specification than the model. As predicted in Chapter 3, the transformation did not stabilize the error variance, confirmed by the results of the Breusch-Pagan test, which reject the homoskedasticity of the residuals at 0.05 significance level. Again, although this is an improvement from the original model, heteroskedasticity is still present.

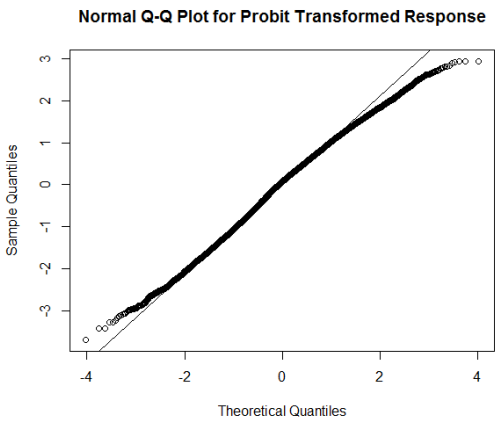

Figure 7.

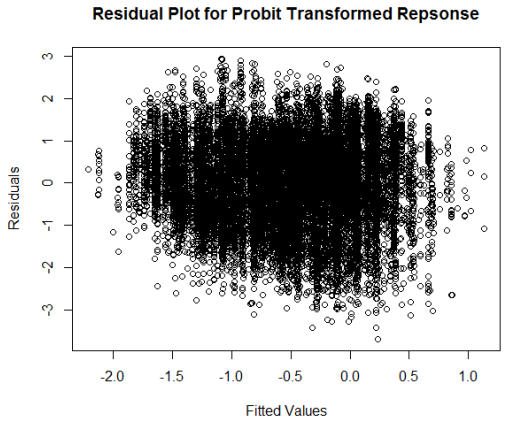

Figure 8.

The second type of transformation of the response variable was a probit transformation. Figure 7 presents the Normal Q-Q Plot for the residuals from the ordinary least squares model in which the response variable has been subject to a probit transformation. Again, the residuals appear much more normal than in the original general linear model. The Shapiro-Wilk test, however, rejects the normality of the residuals at 0.001 significance level. Figure 8 plots the residuals against the fitted values. The residual plot is comparable to that for the logit transformed response. The probit transformation appears to be slightly less preferable to the logit transformation in that the Breusch-Pagan test rejects the homoskedasticity of the residuals for the probit transformation at a much higher significance level.

Considering that all three of these models display heteroskedasticity and non-normal errors, none of the three of the models estimated with ordinary least squares, however, seem very appropriate for proportional data. This conclusion is consistent with what other researchers have found, suggesting that the conditional expectation function is non-linear and that OLS will produce incorrect standard errors (Kieschnick and McCullough 2003; Paolino 2001). The Tobit model is not very appropriate for proportional data either, given that there is no theoretical justification to assume the

dependent variable is normally distributed but can only be observed on (0, 1). Furthermore, even though the censored model is estimated using maximum likelihood, for the interval on which the data are actually observed the regression results will be identical to those of the general linear model, and so all the criticisms that apply to the general linear model for proportional data apply to the Tobit model as well.

These criticisms do not apply to the Fractional Logit model or the Beta regression model. The Fractional logit model explicitly models how the mean proportion relates to the explanatory variables, and Beta regression fully captures the distribution of the dependent variable. The beta distribution also models the non-constant variance, as the variance of a beta-distributed variable is a function of the mean. Akaike's Information Criteria (AIC) is reported in Table 1 for each model in order to provide a formal comparison between the models.[13] The AIC statistics suggest that, of the six models compared in this report, the Beta Regression model is unequivocally the best, followed by the Fractional Logit model. In fact, the probit and logit transformed responses are the least desirable models, even compared to the original general linear model. These findings are consistent with other scholarship on this matter (Kieschnick and McCullough 2003; Paolino 2001). Using simulation techniques, Paolino finds that Beta Regression estimates are less biased and efficient than normal-linear approaches, "especially when the dependent variable has been transformed" (Paolino 2001: 332).

[13] The preferred model is the one with the minimum-valued AIC statistic.

The Effect of Issue, Vote Type, and Routine Matters on Party Conflict Compared Across Models						
	General Linear Model	Logit Transformed Response	Probit Transformed Response	Tobit	Fractional Logit	Beta Regresssion
Type of issue						
Good Government	0.036***	0.281***	0.148***	0.036***	0.165***	0.124***
	(0.007)	(0.051)	(0.027)	(0.007)	(0.036)	(0.029)
Economic	0.119***	0.803***	0.440***	0.119***	0.552***	0.445***
	(0.007)	(0.053)	(0.028)	(0.007)	(0.036)	(0.030)
Social	0.048***	0.279***	0.159***	0.048***	0.228***	0.163***
	(0.006)	(0.041)	(0.022)	(0.006)	(0.025)	(0.023)
Defense	-0.027***	-0.163**	-0.090**	-0.027***	-0.139***	-0.087**
	(0.007)	(0.053)	(0.028)	(0.007)	(0.034)	(0.030)
Foreign	-0.045***	-0.336***	-0.178***	-0.045***	-0.233***	-0.169***
	(0.007)	(0.052)	(0.028)	(0.007)	(0.034)	(0.030)
Type of vote						
Procedural	0.052***	0.280***	0.164***	0.052***	0.223***	0.183***
	(0.005)	(0.039)	(0.020)	(0.005)	(0.025)	(0.022)
Passage	-0.201***	-1.514***	-0.803***	-0.201***	-1.001***	-0.808***
	(0.005)	(0.036)	(0.019)	(0.005)	(0.024)	(0.021)
Routine Matters						
Appropriations	0.030***	0.337***	0.164***	0.030***	0.134***	0.168***
	(0.005)	(0.038)	(0.020)	(0.005)	(0.023)	(0.021)
Purely symbolic	-0.205***	-1.707***	-0.883***	-0.205***	-1.560***	-0.720***
	(0.041)	(0.294)	(0.155)	(0.041)	(0.296)	(0.166)
Presidential Leadership	0.115***	0.893***	0.470***	0.115***	0.574***	0.477***
	(0.006)	(0.042)	(0.022)	(0.006)	(0.025)	(0.024)
Congress						
. . .						
96th Congress	0.034**	0.139	0.089*	0.034**	0.224***	0.027
	(0.011)	(0.079)	(0.042)	(0.011)	(0.049)	(0.045)
97th Congress	0.002	-0.095	-0.038	0.001	0.051	-0.103*
	(0.012)	(0.089)	(0.047)	(0.012)	(0.057)	(0.051)
. . .						
AIC	3120.83	72298.06	49948.02	3159.96	0.94	-10019.55

Standard errors are reported in parentheses below coefficient estimates. *p<.05; **p<.01; ***p<.001

Table 1.

Table 1 presents the coefficient estimates for selected predictor variables and the AIC statistic for each model.[14] The substantive explanatory coefficients are included in the table as well as the results displaying the most discrepancies between models. In spite of the shortcomings of the nearly all of the models, except the Fractional Logit and the Beta Regression models, the inferences made from the coefficient estimates would not differ too much from one model to the next. Absent a direct comparison of the values of the estimates, the inferences made from any of the six models would be relatively similar for all but two coefficients. In order to compare the values of the coefficient estimates directly, the transformation estimates could be mapped back into effects in the 0 to 1 scale as described earlier in Chapter 3. Regardless of model selection, the significance of the coefficient for good government votes indicates that more party conflict occurs on good government issues that on the average issue.

The coefficients for the 96th Congress are statistically significant in all models except for the Beta regression model and the logit transformed response model, and the coefficients for the 97th Congress are never significant except in the Beta regression model. In this model, the indicators for each congress function mostly as a control in order to account for broader factors specific to a given Congress (such as divided or unified government, the majority party's margin of control, etc.). The significance of the coefficients for each Congress does not notably affect the inferences that might be taken from the model, so one might be tempted to dismiss the fact that the significance is not constant for all models. This would be a mistake because in some other model, it could easily be the key explanatory variable of interest that varies in significance across models. The results in this case study confirms what others have shown: some models

[14] See Appendix B for full regression results for each model.

may not pick up on the significance of a regressor and other models may provide a falsely significant estimate, resulting in Type I and Type II errors made in inference (Kieschnick and McCullough 2003). The importance of being discriminating with regard to model choice is clear.

Chapter 5: Conclusion

One drawback of this report is that it only provided one case study. Further research should test these results implementing additional simulations and with a greater diversity of explanatory variables, in order to confirm that these conclusions hold regardless of the type of predictor variables. It may also be important to conduct more rigorous model specification tests as well. Future work might also extend this type of analysis to Bayesian models and time-series models. Indeed, some work on Dynamic Bayesian beta models is already being conducted (da-Silva et al. 2011).

In conclusion, it is clear that model choice matters, especially when modeling proportional data. Although commonly implemented, models that assume a normally distributed response for proportions or attempt to transform proportions to achieve normality are inferior to the alternative approaches. Specifically, normal models can lead to incorrect inferences. Beta regression is unequivocally the best approach for modeling proportions. Fractional Logit is an acceptable alternative as well, although there is evidence that Beta Regression generally outperforms Fractional Logit. It is therefore unclear under what conditions the Fractional Logit might be preferred. If a researcher does wish to use a normal model, they are better off not transforming the dependent variable, as the transformed variable models both performed the worst of all the models considered. Although the censored regression model may take the limit on the dependent variable into account, its results are observationally equivalent to the general linear model. Further, the censored model's assumption that the proportions are merely the observed values of a latent normally distributed variable is no more theoretically sound than the assumption that the proportion is normally distributed. If Beta Regression or

Fractional Logit are not used, the general linear model or the Tobit model will perform about equal to each other and much better than the transformed variable approach.

Appendix A: Notes on Coding

The independent variables in the case study were coded using the Political Institutions and Public Choice House Roll-Call Database, maintained by David Rohde at Duke University, and using the Policy Agendas Project (PAP). The data used from the Policy Agendas Project were originally collected by Frank R. Baumgartner and Bryan D. Jones, with the support of National Science Foundation grant numbers SBR 9320922 and 0111611, and were distributed through the Department of Government at the University of Texas at Austin. Neither NSF nor the original collectors of the data bear any responsibility for the analysis reported here.

Votes coded as "Economic" were those coded into the *Macroeconomics* and *Banking, Finance, and Domestic Commerce* major topic codes in the PAP coding scheme. "Social Issues" were those coded into the *Civil Rights*; *Law, Crime, and Family Issues*; *Health*; *Education*; *Social Welfare*; *Labor, Employment, and Immigration*; and *Community Development and Housing Issues* major topic codes in the PAP coding scheme. "Defense" votes were those coded into the *Defense* major topic code in the PAP coding scheme. "Foreign Issues" were those coded into the *Foreign Trade* and *International Affairs and Foreign Aid* major topic codes in the PAP coding scheme. "Good Government" measures are those that fight corruption, uphold ethical standards, investigate failures, collect and report information, promote fiscal responsibility, ensure electoral integrity, and make government operations more efficient (Lee 2009). Votes coded as "Good Government Issues" were, therefore, those coded into the *Government Efficiency and Bureaucratic Oversight*; *Government Employee Benefits, Civil Service Issues*; *Government Procurement, Procurement Fraud and Contractor Management*; *Presidential Impeachment and Scandal*; *Federal Government Branch Relations and*

Administrative Issues, Congressional Operations; *Regulation of Political Campaigns, Political Advertising, PAC Regulation, Voter Registration, Government Ethics*; and *Census* minor topic codes within the *Government Operations* major topic code in the PAP coding scheme. Votes coded as "Appropriations" were those coded as such in the Rohde/PIPC Roll Call Database. "Purely Symbolic" votes are resolutions that express House or congressional sentiment. To be classified in this category, the resolution cannot require any action from the executive branch or have any other policy content according to the Rohde/PIPC coding scheme.

Appendix B: Results Tables

	β	Std. Error		
Type of issue			F-statistic:	245.7***
Good Government	0.036***	(0.007)		
			N	17492
Economic	0.119***	(0.007)		
			Adj. R^2	0.259
Social	0.048***	(0.006)		
Defense	-0.027***	(0.007)		
Foreign	-0.045***	(0.007)		
Type of vote				
Procedural	0.052***	(0.005)		
Passage	-0.201***	(0.005)		
Routine Matters				
Appropriations	0.030***	(0.005)		
Purely symbolic	-0.205***	(0.041)		
Presidential Leadership	0.115***	(0.006)		
Congress				
94th Congress	0.058***	(0.011)		
95th Congress	0.009	(0.011)		
96th Congress	0.034**	(0.011)		
97th Congress	0.002	(0.012)		
98th Congress	0.092***	(0.012)		
99th Congress	0.147***	(0.012)		
100th Congress	0.158***	(0.012)		
101st Congress	0.120***	(0.012)		
102nd Congress	0.175***	(0.012)		
103rd Congress	0.240***	(0.011)		
104th Congress	0.269***	(0.011)		
105th Congress	0.193***	(0.011)		
106th Congress	0.156***	(0.011)		
107th Congress	0.154***	(0.012)		
108th Congress	0.231***	(0.011)		
			Note:	
Constant	0.257***	(0.009)	*p<.05; **p<.01; ***p<.001	

Table 2. OLS Estimation Results.

	β	Std. Error		
Type of issue			F-statistic:	215.5***
Good Government	0.281***	(0.051)		
			N	17482
Economic	0.803***	(0.053)		
			Adj. R^2	0.235
Social	0.279***	(0.041)		
Defense	-0.163**	(0.053)		
Foreign	-0.336***	(0.052)		
Type of vote				
Procedural	0.280***	(0.039)		
Passage	-1.514***	(0.036)		
Routine Matters				
Appropriations	0.337***	(0.038)		
Purely symbolic	-1.707***	(0.294)		
Presidential Leadership	0.893***	(0.042)		
Congress				
94th Congress	0.354***	(0.079)		
95th Congress	0.038	(0.076)		
96th Congress	0.139	(0.079)		
97th Congress	-0.095	(0.089)		
98th Congress	0.396***	(0.086)		
99th Congress	0.775***	(0.086)		
100th Congress	0.158***	(0.086)		
101st Congress	0.616***	(0.087)		
102nd Congress	0.882***	(0.087)		
103rd Congress	1.276***	(0.082)		
104th Congress	1.588***	(0.079)		
105th Congress	0.956***	(0.081)		
106th Congress	0.642***	(0.081)		
107th Congress	0.628***	(0.085)		
108th Congress	1.098***	(0.080)		
			Note:	
Constant	-1.504***	(0.068)	*p<.05; **p<.01; ***p<.001	

Table 3. OLS Results for Logit Tansformed Response.

	β	Std. Error		
Type of issue			F-statistic:	229.5***
Good Government	0.148***	(0.027)		
			N	17482
Economic	0.440***	(0.028)		
			Adj. R^2	0.246
Social	0.159***	(0.022)		
Defense	-0.090**	(0.028)		
Foreign	-0.178***	(0.028)		
Type of vote				
Procedural	0.164***	(0.020)		
Passage	-0.803***	(0.019)		
Routine Matters				
Appropriations	0.164***	(0.020)		
Purely symbolic	-0.883***	(0.155)		
Presidential Leadership	0.470***	(0.022)		
Congress				
94th Congress	0.200***	(0.042)		
95th Congress	0.023	(0.040)		
96th Congress	0.089*	(0.042)		
97th Congress	-0.038	(0.047)		
98th Congress	0.248***	(0.046)		
99th Congress	0.452***	(0.046)		
100th Congress	0.465***	(0.045)		
101st Congress	0.365***	(0.045)		
102nd Congress	0.527***	(0.046)		
103rd Congress	0.748***	(0.044)		
104th Congress	0.901***	(0.042)		
105th Congress	0.576***	(0.043)		
106th Congress	0.413***	(0.042)		
107th Congress	0.0406***	(0.045)		
108th Congress	0.675***	(0.042)		
			Note:	
Constant	-0.843***	(0.036)	*p<.05; **p<.01; ***p<.001	

Table 4. OLS Results for Probit Transformed Response.

	β	Std. Error		
Type of issue			Chi-squared (25):	5268.8***
Good Government	0.036***	(0.007)		
			N	17482
Economic	0.119***	(0.007)		
			Pseudo. R^2	0.6291
Social	0.048***	(0.006)		
			Log-Likelihood:	-1552.98
Defense	-0.027***	(0.007)		
			Log σ:	-1.331***
Foreign	-0.045***	(0.007)		(-0.005)
Type of vote				
Procedural	0.052***	(0.005)		
Passage	-0.201***	(0.005)		
Routine Matters				
Appropriations	0.030***	(0.005)		
Purely symbolic	-0.205***	(0.041)		
Presidential Leadership	0.115***	(0.006)		
Congress				
94th Congress	0.058***	(0.011)		
95th Congress	0.009	(0.011)		
96th Congress	0.034**	(0.011)		
97th Congress	0.002	(0.012)		
98th Congress	0.092***	(0.012)		
99th Congress	0.147***	(0.012)		
100th Congress	0.158***	(0.012)		
101st Congress	0.120***	(0.012)		
102nd Congress	0.175***	(0.012)		
103rd Congress	0.240***	(0.011)		
104th Congress	0.269***	(0.011)		
105th Congress	0.193***	(0.011)		
106th Congress	0.156***	(0.011)		
107th Congress	0.0154***	(0.012)		
108th Congress	0.231***	(0.011)		
			Note:	
Constant	0.257***	0.009	*p<.05; **p<.01; ***p<.001	

Table 5. MLE Results for Tobit Model.

	β	Std. Error		
Type of issue			Log	
Good Government	0.165***	(0.036)	pseudolikelihood:	-8184.2
Economic	0.552***	(0.036)	N	17492
Social	0.228***	(0.025)	Deviance	6223.3
Defense	-0.139***	(0.034)	Pearson	5652.0
Foreign	-0.233***	(0.034)		
Type of vote				
Procedural	0.223***	(0.025)		
Passage	-1.001***	(0.024)		
Routine Matters				
Appropriations	0.134***	(0.023)		
Purely symbolic	-1.560***	(0.296)		
Presidential Leadership	0.574***	(0.025)		
Congress				
94th Congress	0.335***	(0.046)		
95th Congress	0.0699	(0.046)		
96th Congress	0.224***	(0.049)		
97th Congress	0.051	(0.057)		
98th Congress	0.517***	(0.053)		
99th Congress	0.766***	(0.050)		
100th Congress	0.822***	(0.050)		
101st Congress	0.643***	(0.051)		
102nd Congress	0.887***	(0.052)		
103rd Congress	1.158***	(0.051)		
104th Congress	1.285***	(0.048)		
105th Congress	0.972***	(0.053)		
106th Congress	0.812***	(0.056)		
107th Congress	0.807***	(0.057)		
108th Congress	1.158***	(0.052)		
			Note:	
Constant	-1.158***	(0.040)	*p<.05; **p<.01; ***p<.001	

Table 6. QMLE Results for Fractional Logit.

	μ	Std. Error		
Type of issue			Chi-suqared (25):	4295.5***
Good Government	0.124***	(0.029)		
			N	17482
Economic	0.445***	(0.030)		
			Pseudo R^2	0.235
Social	0.163***	(0.023)		
			Log Likelihood:	5037
Defense	-0.087**	(0.030)		
			Phi	2.095***
Foreign	-0.169***	(0.030)		(0.020)
Type of vote				
Procedural	0.183***	(0.022)		
Passage	-0.808***	(0.021)		
Routine Matters				
Appropriations	0.168***	(0.021)		
Purely symbolic	-0.720***	(0.166)		
Presidential Leadership	0.477***	(0.024)		
Congress				
94th Congress	0.187***	(0.045)		
95th Congress	-.005	(0.043)		
96th Congress	0.027	(0.045)		
97th Congress	-0.103*	(0.051)		
98th Congress	0.184***	(0.049)		
99th Congress	0.420***	(0.050)		
100th Congress	0.447***	(0.049)		
101st Congress	0.340***	(0.050)		
102nd Congress	0.497***	(0.049)		
103rd Congress	0.792***	(0.047)		
104th Congress	0.915***	(0.045)		
105th Congress	0.547***	(0.046)		
106th Congress	0.394***	(0.046)		
107th Congress	0.0409***	(0.048)		
108th Congress	0.669***	(0.046)		
			Note:	
Constant	-0.859***	0.039	*p<.05; **p<.01; ***p<.001	

Table 8. MLE Results for Beta Regression.

References

Aitchison, J. 1986. *The Statistical Analysis of Compositional Data*. New York: Chapman and Hall.

Baumgartner, Frank R., and Bryan D. Jones. *Policy Agendas Project* (1973-2004), NSF SBR9320922 and 0111611.

Buis, Maarten. 2006. "Proportions as a Dependent Variable." In *12th UK Stata Users Group Meeting*, London.

———. 2010. "Analyzing Proportions." In *8th German Stata Users Goup Meeting*, Berlin, Germany.

Da-Silva, C.Q., H.S. Migon, L.T. Correia. 2011. "Dynamic Bayesian beta models." *Computational Statistics and Data Analysis* 55: 2074-2089.

Ferrari, Silvia L.P., and Francisco Cribari-Neto. 2004. "Beta Regression for Modelling Rates and Proportions." *Journal of Applied Statsitics* 31: 799–815.

Fox, John. 2008. *Applied Regression Analysis and Generalized Linear Models*. Second Edi. Los Angeles: Sage Publications.

Greene, William. 2008. *Econometric Analysis*. Sixth Edit. Upper Saddle River, NJ: Pearson.

Hox, Joop. 2010. *Multilevel Analysis: Techniques and Applications*. Second Edi. New York: Routledge.

Kieschnick, Robert, and BD McCullough. 2003. "Regression Analysis of Variates Observed on (0, 1): Percentages, Proportions and Fractions." *Statistical Modelling* 3(3): 193–213.

Kmenta, Jan. 1997. *Elements of Econometrics*. Second Edi. Ann Arbor: The University of Michigan Press.

Lee, Frances E. 2009. *Beyond Ideology: Politics, Principles, and Partisanship in the U.S. Senate*. Chicago: University of Chicago Press.

Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage Publications.

"NIST/SEMATECH e-Handbook of Statistical Methods." http://www.itl.nist.gov/div898/handbook/ (August 5, 2013).

Paolino, Philip. 2001. "Maximum Likelihood Estimation of Models with Beta-Distributed Dependent Variables." *Political Analysis* 9(4): 325–46.

Papke, Leslie E., and Jeffrey M. Wooldridge. 1996. "Econometric Methods for Fractional Response Variables With an Application to 401(k) Plan Participation Rates." *Journal of Applied Econometrics* 11(6): 619–32.

Rice, Stuart A. 1925. "The Behavior of Legislative Groups: A Method of Measurement." *Political Science Quarterly* 40(1): 60–72.

Rohde, David. 2010. Political Institutions and Public Choice House Roll Call Dataset. Duke University, Durham, NC.

Stokes, Donald E. 1963. "Spatial Models of Party Competition." *American Political Science Review* 57(2): 368-377.

Theriault, Sean M. 2008. *Party Polarization in Congress*. New York: Cambridge University Press.